

【特許請求の範囲】

【請求項1】 利用者からのサービス要求を通信手段を介して受信および制御を行なうサービス制御部と、サービスを実現するためのアプリケーションの実行および管理を行なう複数のアプリケーション実行管理部と、ストリームデータを蓄積し利用者に対してストリームの送出を行なう複数のストリーム送出部と、前記複数のストリーム送出部を制御するストリーム制御部と、前記ストリーム送出部に蓄積されているストリームデータの管理情報や詳細情報を蓄積しているコンテンツデータ管理部とからなる、通信手段を介して接続されている利用者端末に対して各種サービスを提供するインタラクティブTVシステムにおけるアプリケーションサーバシステムであって、アプリケーションの起動や停止などのアプリケーションの実行を制御する前記アプリケーション実行管理部は、アプリケーションの実行状態を記憶するアプリケーション実行状態情報と、そのアプリケーションを構成するプロセスの実行状態を記憶するプロセス実行状態情報と、アプリケーションを利用している利用者の利用状況を記憶するアプリケーション利用状況情報と、動作するサーバの負荷状態を記憶する負荷状態情報とを少なくとも所有し、前記アプリケーション実行状態情報と、前記プロセス実行状態情報と、前記アプリケーション利用状況情報と、前記負荷状態情報に基づき、前記複数のアプリケーション実行管理部間の負荷分散を行なうことを特徴とするアプリケーションサーバシステム。

【請求項2】 前記複数のアプリケーション実行管理部は、当該複数のアプリケーション実行管理部を制御するプライマリアプリケーション実行管理部と、前記プライマリアプリケーション実行管理部において障害が発生した場合に、前記プライマリアプリケーション実行管理部を代替するセカンダリアプリケーション実行管理部と、前記プライマリアプリケーション実行管理部および前記セカンダリアプリケーション実行管理部以外の複数の一般アプリケーション実行管理部とにより構成され、前記プライマリアプリケーション実行管理部は、当該プライマリアプリケーション実行管理部を含む全ての前記複数のアプリケーション実行管理部の構成情報を所有し、前記複数の一般アプリケーション実行管理部内の各々の一般アプリケーション実行管理部は、当該複数の一般アプリケーション実行管理部間において、重なることのない一つの順位を有することを特徴とする請求項1記載のアプリケーションサーバシステム。

【請求項3】 前記複数のアプリケーション実行管理部は、障害対策を実施する手段を更に具備し、当該障害対策を実施する手段は、

前記セカンダリアプリケーション実行管理部と、前記複数の一般アプリケーション実行管理部が、前記プライマリアプリケーション実行管理部に対して、予め定めた障害認定時間より短い時間間隔で生存通知メッセージを送る

手段と、

前記プライマリアプリケーション実行管理部が、前記セカンダリアプリケーション実行管理部に対して、前記障害認定時間より短い時間間隔で生存通知メッセージを送る手段と、

前記プライマリアプリケーション実行管理部が、前記複数の一般アプリケーション実行管理部より、前記障害認定時間を越えても生存通知メッセージが到着しない場合、システムから当該一般アプリケーション実行管理部を切断する手段と、

10 前記プライマリアプリケーション実行管理部が、前記セカンダリアプリケーション実行部より、前記障害認定時間を越えても生存通知メッセージが到着しない場合、前記複数の一般アプリケーション実行管理部の中から、前記順位の最も高い一般アプリケーション実行管理部を選択し、当該一般アプリケーション実行管理部に対して前記セカンダリアプリケーション実行管理部にトランスフォームするためのトランスフォームメッセージを送る手段と、

20 前記セカンダリアプリケーション実行管理部が、前記プライマリアプリケーション実行管理部より、前記障害認定時間を越えても生存通知メッセージが到着しない場合、当該セカンダリアプリケーション実行管理部が、プライマリアプリケーション実行管理部にトランスフォームするとともに、前記複数の一般アプリケーション実行管理部の中から、前記順位の最も高い一般アプリケーション実行管理部を選択し、当該一般アプリケーション実行管理部に対して前記セカンダリアプリケーション実行管理部にトランスフォームするためのトランスフォームメッセージを送る手段から構成されることを特徴とする請求項2記載のアプリケーションサーバシステム。

【請求項4】 前記プライマリアプリケーション実行管理部は、前記セカンダリアプリケーション実行管理部と前記複数の一般アプリケーション実行管理部の負荷状態を当該プライマリアプリケーション実行管理部の負荷状態情報に記憶する手段と、前記アプリケーション実行状態情報に、アプリケーション毎にアプリケーションの負荷特性を示す負荷特性情報を記憶する手段を具備し、

さらに、前記プライマリアプリケーション実行管理部におけるアプリケーションの起動において、当該アプリケーションに対応する前記負荷特性情報を参照するとともに前記負荷状態情報を検索し、前記負荷特性に最も適合したアプリケーション実行管理部を選択する手段を具備して、負荷分散を実現することを特徴とする請求項2記載のアプリケーションサーバシステム。

【請求項5】 前記プライマリアプリケーション実行管理部は、前記アプリケーション実行状態情報に、アプリケーション毎にアプリケーションの起動特性を示す起動特性情報を記憶する手段を具備するとともに、前記プライマリアプリケーション実行管理部におけるアプリケー

ションの起動において、当該アプリケーションに対応する前記起動特性情報により、当該アプリケーションの起動を制御する手段を具備して、アプリケーションの起動を行なうことを特徴とする請求項2記載のアプリケーションサーバシステム。

【請求項6】 前記複数のアプリケーション実行管理部は、構成管理手段を更に具備し、当該構成管理手段は、アプリケーション実行管理部の新規接続において、当該アプリケーション実行管理部が、前記プライマリアプリケーション実行管理部に対して接続要求メッセージを送る手段と、

前記プライマリアプリケーション実行管理部が、当該アプリケーション実行管理部からの接続要求メッセージを受けて、当該アプリケーション実行管理部に対して最新の前記アプリケーション実行状態情報と、プロセス実行状態情報と、アプリケーション利用状況情報と、負荷状態情報を転送する手段を具備して、構成管理を行なうことを特徴とする請求項2記載のアプリケーションサーバシステム。

【請求項7】 前記サービス制御部は、アプリケーションに対して、自制御部と前記アプリケーション実行管理部と前記ストリーム制御部と前記コンテンツデータ管理部の各部の機能を、アプリケーションプログラム・インタフェースライブラリとして提供することを特徴とする請求項1記載のアプリケーションサーバシステム。

【請求項8】 前記サービス制御部が、オブジェクト分散システムであって、サービスを分散オブジェクトとして管理することを特徴とする請求項1記載のアプリケーションサーバシステム。

【請求項9】 前記コンテンツデータ管理部が、オブジェクトデータベースシステムであって、前記ストリームデータの管理情報、詳細情報をオブジェクトとして管理することを特徴とする請求項1記載のアプリケーションサーバシステム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、ネットワーク環境下において、動画の再生、送出をともなうインタラクティブなアプリケーションサービスを実行、管理するために有用なアプリケーションサーバシステムに関し、特に、どのような負荷特性を持つアプリケーションが混在していても、負荷分散が可能なアプリケーションサーバシステムに関するものである。

【0002】

【従来の技術】近年、MPEGやATMなどの動画の圧縮技術や、大容量データ通信技術の進展にもない、それらの技術に応じたインタラクティブTVシステムが具現化され始めている。とりわけ利用者の要求に対して実時間の保証が必須である、動画の再生や送出をともなうアプリケーションサービスの実行、管理が重要視

されている。

【0003】以下、このような動画の再生や送出をともなうアプリケーションサービスの実行、管理を行なうアプリケーションサーバシステムを説明する。

【0004】従来、動画の再生や送出をともなうアプリケーションサービスの実行は、専用のビデオサーバシステムと、汎用の計算機システム上に構築されたアプリケーションシステムを結合することにより行なわれていた。専用のビデオサーバシステムを用いない場合は、汎用の計算機上に構築されたアプリケーションシステムに、動画の再生や送出を行なう機能を付加することにより実現している。

【0005】利用者に対してサービスを提供するアプリケーションの起動および停止は、アプリケーションシステムの立上り時に起動し、アプリケーションシステムの停止時に停止するか、利用者からの要求が到着した時に起動し、サービスの終了時に停止するかのいずれかである。

【0006】利用者に対してサービスを提供するアプリケーションは、利用者から要求を受け付け、動画の再生や送出を行なう。利用者数の増大でアプリケーションの負荷が増加した場合は、専用ビデオサーバシステムを用いていはいはビデオサーバシステムを増設することによって負荷の分散を図れる。また、汎用の計算機上に構築されたシステムにおいては、プロセッサの多重化や処理性能の高い計算機への移行によって、負荷分散や処理性能の向上を図っている。

【0007】アプリケーションシステムの障害対策においては、フォールトトレラント機などの高信頼性の計算機を用いることによって実現している。

【0008】

【発明が解決しようとする課題】しかしながら、現在、動画の再生や送出をともなうサービスの増加にともない、同じアプリケーションサーバシステムにおいて、異なる負荷特性のサービスの混在が要求されるようになってきた。異なる負荷特性のサービスとは、例えば、ビデオ・オン・デマンドサービス等の、動画の再生や送出が中心となるサービスや、TVショッピングサービスやマルチメディア情報検索サービス等の、インタラクティブな操作が中心となるサービス等である。

【0009】従来の技術では、動画の再生や送出をともなうアプリケーションサービスの実行は、専用のビデオサーバシステムと、汎用の計算機システム上に構築されたアプリケーションシステムの結合、もしくは汎用の計算機上に構築されたアプリケーションシステムに、動画の再生や送出を行なう機能を付加することによって実現しているため、動画の再生や送出に関する負荷状態の管理や、動画の再生や送出にともなう負荷とアプリケーションの負荷との分離が困難であった。

【0010】また、負荷状態の判定においても、CPU

やネットワーク装置、動画像の再生装置や送出装置等のハードウェアの負荷状態の監視が主であって、アプリケーションが固有に持っている負荷特性に基づく負荷状態の判定や管理は困難であった。

【0011】さらに、利用者に対してサービスを提供するアプリケーションの起動および停止においても、アプリケーションシステムの立上げ時に起動し、アプリケーションシステムの停止時に停止するのでは、CPUやメモリなどの計算機上の資源の効率的な使用が困難であり、また、利用者から要求が到着した時に起動し、サービスの終了時に停止させるのでは、利用者の要求からサービスの開始までの応答時間の短縮が困難であった。

【0012】さらにまた、従来のシステムにおいては、アプリケーションサーバシステムの障害対策の実現においては、フォールトトレラント計算機の導入や、アプリケーション自体に障害管理機能を実装することで信頼性を向上させている。このため、障害対策のためのシステム構築コストが多く必要であった。

【0013】このように、従来のアプリケーションシステムでは、専用のビデオサーバシステムか、汎用の計算機上で構築されたアプリケーションシステムによって構築されていたため、多様なアプリケーションサービスのための負荷分散や、柔軟な障害対策を実現することが困難であった。

【0014】前記のようなアプリケーションサーバシステムの困難性を克服するために本発明は、動画像の再生や送出を行なうアプリケーションに対し、汎用的な動作環境を提供するものであって、どのような負荷特性を持つアプリケーションが混在していても、負荷分散が可能なアプリケーションサーバシステムを実現し、かつ負荷分散とともに、アプリケーションサーバシステムの障害管理をも可能とするアプリケーションサーバシステムを実現することを目的とする。

【0015】

【課題を解決するための手段】本発明は、前記目的を達成するために、ストリーム制御部と、アプリケーション制御を行なうアプリケーション実行管理部を分離することにより、それぞれの負荷状態を明確にするとともに、負荷状態の判定において、アプリケーションサービス登録時に負荷判定のための条件を指定することにより、多様な負荷特性を持つアプリケーションサーバが混在する様な場合でも最適な負荷分散を行なえるようにしている。

【0016】また、アプリケーションの起動特性に基づき、アプリケーションの起動および終了を制御することにより計算機上の資源の効率的な利用を行なえるようにしている。

【0017】さらに、アプリケーション実行管理部に障害管理機能を持たせ、監視されているアプリケーション実行管理部が動作するサーバの状態をプライマリアプリ

ケーション実行管理部に通知することにより、自律分散的な障害管理を行なうようにしている。

【0018】

【発明の実施の形態】以下、本発明の実施の形態を、図面を参照して説明する。図1は、本発明の実施の形態におけるアプリケーションサーバシステムのシステム概要図である。

【0019】図1において、アプリケーションサーバシステムは、通信手段105を介して利用者端末106からのサービス要求を処理するサービス制御部100と、サービス制御部100からサービスを実現するためのアプリケーションの起動要求を受け付け、アプリケーションの実行や制御を行なうアプリケーション実行管理部101と、アプリケーション実行管理部101によって起動されたアプリケーションからのストリーム送出要求や、利用者端末106からのストリーム送出要求等の、ストリームに関する制御を行なうストリーム制御部103と、利用者端末106に対してストリームの送出を行なうストリーム送出部104と、ストリーム送出部104によって送出されるストリームの詳細情報を管理しているコンテンツ情報管理部102から構成されている。ストリーム制御部103は、ストリームに関する制御の他に、ストリーム送出部104の制御を行なう。

【0020】次に、アプリケーション実行管理部101内におけるアプリケーション実行状態情報の論理構造について説明する。まず、アプリケーション実行状態情報120は、図2に示す構造により格納されている。アプリケーションID121は、アプリケーションを特定するための一意の番号である。

【0021】起動特性122は、アプリケーションの起動および停止時の処理の特性を示している。利用者数123は、アプリケーションを利用している利用者の数を示している。プロセス数124は、アプリケーション実行管理部101内において、アプリケーションのために起動しているプロセスの数を示している。最大プロセス数125は、アプリケーション実行管理部101内において、アプリケーションのために起動することができるプロセス数の上限値を示している。

【0022】最大利用者数126は、1プロセスに対して接続することができる利用者数の上限値を示している。負荷特性127は、アプリケーションの負荷特性を示している。プロセス情報ポインタ128は、起動したプロセスのプロセス実行状態情報を参照するためのポインタである。

【0023】次に、アプリケーション実行管理部101内におけるプロセス実行状態情報142の論理構造について説明する。まず、プロセス実行状態情報142は、図3に示す構造により格納されている。

【0024】アプリケーション実行状態情報120内のプロセス情報ポインタ128は、プロセス実行状態情報142リ

7

ストの先頭を指す先頭ポインタ140と、最後を指す最後ポインタ141により構成されている。プロセス実行状態情報142内の次ポインタ143と前ポインタ144は、リスト構造を実現するためのプロセス実行状態情報ポインタである。プロセスID145は、アプリケーション実行管理部101において、プロセスを一意に特定するための番号である。状態146は、プロセスの状態を示している。

【0025】利用者数147は、プロセスに接続している利用者の数を示している。利用者情報ポインタ148は、接続している利用者のアプリケーション利用状況情報を参照するためのポインタである。アプリケーション情報ポインタ149は、対応するアプリケーションのアプリケーション実行状態情報120を参照するためのポインタである。

【0026】次に、アプリケーション実行管理部101内における利用状況情報の論理構造について説明する。まず、アプリケーション利用状況情報162は、図4に示す構造により格納されている。

【0027】プロセス実行状態情報142内の利用者情報ポインタ148は、アプリケーション利用状況情報162のリストの先頭を指す先頭ポインタ160と、リストの最後を指す最後ポインタ161により構成されている。アプリケーション利用状況情報162内の次ポインタ163と前ポインタ164は、リスト構造を実現するためのアプリケーション利用状況情報参照ポインタである。

【0028】接続パラメータ165は、アプリケーション接続時に指定したパラメータを記憶している。利用者ID166はアプリケーションサーバシステム内において利用者を一意に特定するためのIDである。スレッドID167は、対応するプロセス内において、利用者からのメッセージを待ち受けるためのスレッドを一意に特定するための番号である。接続時刻168は、プロセスへの接続時刻を示している。プロセス情報ポインタ169は、対応するプロセスの、プロセス実行状態情報を参照するためのポインタである。

【0029】次に、本実施の形態の動作の概略について説明する。まず、アプリケーションサーバシステムは、図5に示すように、複数の一般アプリケーション実行管理部182と、複数の一般アプリケーション実行管理部182を制御するプライマリアプリケーション実行管理部180と、プライマリアプリケーション実行管理部180において障害が発生した場合にプライマリアプリケーション実行管理部180を代替するセカンダリアプリケーション実行管理部181とを備えている。

【0030】また、図6に示すように、アプリケーション実行管理部200は、プライマリアプリケーション実行管理部180、セカンダリアプリケーション実行管理部181、一般アプリケーション実行管理部182の区別なく、アプリケーション205の起動や停止など、アプリケーション205の実行を制御するものであって、アプリケーション

8

ンの状態を記憶するアプリケーション実行状態情報201と、アプリケーション205を構成するプロセスの実行状態を記憶するプロセス実行状態情報202と、アプリケーションを利用している利用者の利用状況を記憶するアプリケーション利用状況情報203と、アプリケーション実行管理部が動作するサーバの負荷状態を記憶する負荷状態情報204を保持している。

【0031】プライマリアプリケーション実行管理部180の負荷状態情報220は、図7に示す構造により格納されている。負荷状態情報220は、プライマリアプリケーション実行管理部180の固有の負荷情報221の他に、セカンダリアプリケーション実行管理部181固有の負荷情報222および一般アプリケーション実行管理部182固有の負荷情報223を管理している。セカンダリアプリケーション実行管理部181は、プライマリアプリケーション実行管理部180の負荷状態情報220の複写を保持している。

【0032】アプリケーション実行管理部の負荷分散について、図8を参照して説明する。プライマリアプリケーション実行管理部180は、サービス制御部100からのアプリケーション起動要求を受け付けると（ステップ240）、アプリケーション実行状態情報120を参照し（ステップ241）、更にアプリケーション実行状態情報120中の負荷特性127を参照する（ステップ242）（関連の図2参照）。

【0033】次に、負荷特性127をもとに選択条件テーブル243を参照し（ステップ244）、負荷特性127に最も適したアプリケーション実行管理部101を選択するための選択条件を取得する（ステップ245）。

【0034】次に、その選択条件を評価するための、アプリケーション実行管理部選択ルーティーン247を実行する（ステップ246）。アプリケーション実行管理部選択ルーティーン247は、負荷状態情報220（図7も参照）を検索し（ステップ248）、選択条件を評価し、最適なアプリケーション実行管理部101を選択し、アプリケーションの起動を要求する（ステップ249）。

【0035】これにより、負荷特性の異なるアプリケーションの混在においても、アプリケーション毎に最も適したアプリケーション実行管理部101を選択することができる。

【0036】次に、アプリケーションの起動について、図9の流れ図を参照して説明する。まず、利用者端末106からのサービス要求は、通信手段105を介してサービス制御部100に伝えられる。サービス制御部100は、プライマリアプリケーション実行管理部180に対してアプリケーションの起動要求を行なう。プライマリアプリケーション実行管理部180は、負荷特性127に基づき、最も適したアプリケーション実行管理部を選択し、アプリケーションの起動要求を行なう（ステップ260）（関連の図1及び図8参照）。

【0037】起動要求を受けた（ステップ261）アプリ

ケーション実行管理部は、アプリケーション実行状態情報120内の起動特性122（図2参照）を参照し（ステップ262）、起動特性122がマルチユーザ型でなければ、アプリケーションを起動し（ステップ268）、起動したアプリケーションと利用者端末106を接続する（ステップ267）。

【0038】起動特性122がマルチユーザ型であれば、アプリケーション実行状態情報120内のプロセス情報ポイント128（図2参照）を参照し、利用者数123が最大利用者数126未満であるプロセス実行状態情報142（図2参照）を検索する（ステップ263）。利用者数123が最大利用者数126未満であるプロセス実行状態情報142が存在したら、プロセス実行状態情報142に対応するアプリケーションと利用者端末106を接続する（ステップ267）。

【0039】プロセス実行状態情報142が存在しなければ、アプリケーション実行状態情報120内の最大プロセス数125とプロセス数124（図2参照）を比較して（ステップ265）、プロセス数124が最大プロセス数125を越えていなければ、アプリケーションを起動し（ステップ266）、起動したアプリケーションと利用者端末106を接続する（ステップ267）。プロセス数124が最大プロセス数125を越えていなければ、プライマリアプリケーション実行管理部が、次に最も適したアプリケーション実行管理部を選択し、再度アプリケーション起動要求を行なう（ステップ269）。

【0040】次に、アプリケーションの停止について図10の流れ図を参照して述べる。まず、利用者端末106からのサービス停止要求は、通信手段105によりアプリケーション実行管理部101内で実行しているアプリケーションに伝えられる。

【0041】アプリケーションは、アプリケーション実行管理部101に対して、アプリケーションの停止を要求する（ステップ280）。アプリケーション実行管理部101は、アプリケーション実行状態情報120内の起動特性122（図2参照）を参照し（ステップ281）、起動特性122が常駐型であれば、アプリケーションと利用者端末106とを切断する（ステップ285）。

【0042】もし、起動特性122がシングルユーザ型であれば（ステップ282）、アプリケーションの停止を行なう（ステップ284）。もし、起動特性122がマルチユーザ型であり、利用者が一人以下であれば（ステップ283）、アプリケーションを停止する（ステップ284）。もしマルチユーザ型で利用者が2人以上であれば、アプリケーションと利用者端末106とを切断する（ステップ285）。

【0043】このようにして、起動特性122によってアプリケーションの起動を制御することにより、CPUやメモリなどのシステム資源を効率的に使用することができる。

【0044】次に、アプリケーション実行管理部の障害

管理について説明する。

【0045】まず、一般アプリケーション実行管理部182の障害管理について、図11を参照して説明する。一般アプリケーション実行管理部182は、障害認定時間未満の時間間隔でプライマリアプリケーション実行管理部180に対して生存通知メッセージ300を送っている。プライマリアプリケーション実行管理部180は、一般アプリケーション実行管理部182からの生存通知メッセージ300の到着時間の間隔を監視している。

【0046】プライマリアプリケーション実行管理部180は、生存通知メッセージ300が障害認定時間を越えても到着しない場合、一般アプリケーション実行管理部182に障害が発生したと認定し、セカンダリアプリケーション実行管理部181に対して、当該一般アプリケーション実行管理部182の切断を要求する切り離しメッセージ301を送る。

【0047】セカンダリアプリケーション実行管理部181は、切り離しメッセージ301の受信後、当該一般アプリケーション実行管理部182の切断を行ない（ステップ302）、プライマリアプリケーション実行管理部180に対して、切り離し完了メッセージ303を送る。プライマリアプリケーション実行管理部180は、セカンダリアプリケーション実行管理部181からの切り離し完了メッセージ303の受信後、当該一般アプリケーション実行管理部182の切断を行なう（ステップ304）。

【0048】次に、セカンダリアプリケーション実行管理部の障害管理について、図12を参照して説明する。セカンダリアプリケーション実行管理部181は、障害認定時間未満の時間間隔でプライマリアプリケーション実行管理部180に対して生存通知メッセージ320を送っている。プライマリアプリケーション実行管理部180は、セカンダリアプリケーション実行管理部181からの生存通知メッセージ320の到着時間の間隔を監視している。

【0049】プライマリアプリケーション実行管理部180は、生存通知メッセージ320が障害認定時間を越えても到着しない場合、セカンダリアプリケーション実行管理部181に障害が発生したと認定し、複数の一般アプリケーション実行管理部182の中から最も順位の高い一般アプリケーション実行管理部182を選択し（ステップ321）、当該一般アプリケーション実行管理部182に対して、セカンダリアプリケーション実行管理部181へのトランスフォームを要求する、トランスフォームメッセージ322を送る。

【0050】当該一般アプリケーション実行管理部182は、トランスフォームメッセージ322の受信後、セカンダリアプリケーション実行管理部181へトランスフォームし（ステップ323）、プライマリアプリケーション実行管理部180に対してトランスフォーム完了メッセージ324を送る。

【0051】次に、プライマリアプリケーション実行管

11

理部の障害管理について、図13を参照して説明する。プライマリアプリケーション実行管理部180は、障害認定時間未満の時間間隔でセカンダリアプリケーション実行管理部181に対して生存通知メッセージ340を送っている。セカンダリアプリケーション実行管理部181は、プライマリアプリケーション実行管理部180からの生存通知メッセージ340の到着時間の間隔を監視している。

【0052】セカンダリアプリケーション実行管理部181は、生存通知メッセージ340が障害認定時間を越えても到着しない場合、プライマリアプリケーション実行管理部180に障害が発生したと認識し、プライマリアプリケーション実行管理部180へトランスフォームする(ステップ341)。

【0053】新プライマリアプリケーション実行管理部180は、サービス制御部100に対してプライマリアプリケーション実行管理部代替メッセージ342を送るとともに、複数の一般アプリケーション実行管理部182の中から最も順位の高い一般アプリケーション実行管理部182を選択して(ステップ343)、当該一般アプリケーション実行管理部182に対して、セカンダリアプリケーション実行管理部181へのトランスフォームを要求する、トランスフォームメッセージ344を送る。

【0054】当該一般アプリケーション実行管理部182は、トランスフォームメッセージ344の受信後、セカンダリアプリケーション実行管理部181へトランスフォームし(ステップ345)、プライマリアプリケーション実行管理部180に対してトランスフォーム完了メッセージ46を送る。プライマリアプリケーション実行管理部180は、サービス制御部100に対してトランスフォーム完了メッセージ347を送る。

【0055】次に、一般アプリケーション実行管理部182の構成管理について、図14を参照して説明する。一般アプリケーション実行管理部182を新規接続する場合、一般アプリケーション実行管理部182は、プライマリアプリケーション実行管理部180に対して接続要求メッセージ360を送る。プライマリアプリケーション実行管理部180は、一般アプリケーション実行管理部182からの接続要求メッセージ360を受信すると、最新のアプリケーション実行状態情報361と、プロセス実行状態情報362と、アプリケーション利用状況情報363と、負荷状態情報364を、接続要求メッセージ360を送信してきた一般アプリケーション実行管理部182に送る。

【0056】これらの情報を受信した後に、新規接続要求をした一般アプリケーション実行管理部182は、接続が可能であるかを判断し、可能であれば一般アプリケーション実行管理部の一員となることができる。

【0057】以上では、アプリケーションサーバシステムのアプリケーション実行管理部について説明してきたが、ここでは、サービス制御部およびコンテンツ情報(データ)管理部について補足説明する。

12

【0058】利用者端末106からのサービス要求を処理するサービス制御部100は、サービスを実現するためのアプリケーションの起動、実行及びストリームに関する制御、データの管理などの並びに自身の制御などの各部の機能のため、アプリケーションに対してアプリケーション・インタフェースライブラリを保持しており、各部の要求に応じてアプリケーション・インタフェースライブラリを提供することができる。

【0059】また、サービス制御部100は、サービスを分散オブジェクトとして管理できるようにするために、オブジェクト分散システムとすることができる。

【0060】さらに、コンテンツ情報(データ)管理部102は、ストリームデータの管理情報や詳細情報をオブジェクトとして管理するために、オブジェクトデータベースシステムとして構成することができる。

【0061】

【発明の効果】以上のように、本発明によれば、ストリーム制御部と、アプリケーション制御を行なうアプリケーション実行管理部を分離させているため、それぞれの負荷状態を明確にすることが可能となる。

【0062】また、負荷状態の判定においても、アプリケーションサービス登録時に負荷特性を指定することにより、多様な負荷特性を持つアプリケーションが混在する様な場合であっても、最適な負荷分散を行なうことが可能である。

【0063】また、アプリケーションの起動特性に基づき、アプリケーションの起動および停止を制御することで、計算機上のシステム資源を効率的に利用することが可能となる。

【0064】さらに、アプリケーション実行管理部に障害管理機能を持たせ、監視されているアプリケーション実行管理部が動作するサーバの状態をプライマリアプリケーション実行管理部に通知することで、自律分散的な障害管理を行うことが可能である。

【図面の簡単な説明】

【図1】本発明に係るアプリケーションサーバシステムの実施の形態を示すシステム構成図、

【図2】本発明に係るアプリケーション実行状態情報の論理構造を示す図、

【図3】本発明に係るプロセス実行状態情報の論理構造を示す図、

【図4】本発明に係るアプリケーション利用状況情報の論理構造を示す図、

【図5】本発明に係るアプリケーション実行管理部を示す図、

【図6】本発明に係るアプリケーション実行管理部内部の情報を示す図、

【図7】本発明に係るプライマリアプリケーション実行管理部の負荷状態情報の論理構造を示す図、

【図8】本発明に係るプライマリアプリケーション実行

13

管理部による負荷分散を表す図、

【図9】本発明に係るアプリケーションの起動手順を示すブロック図、

【図10】本発明に係るアプリケーションの停止手順を示すブロック図、

【図11】本発明に係る一般アプリケーション実行管理部の障害管理を示すシーケンス図、

【図12】本発明に係るセカンダリアプリケーション実行管理部の障害管理を示すシーケンス図、

【図13】本発明に係るプライマリアプリケーション実行管理部の障害管理を示すシーケンス図、

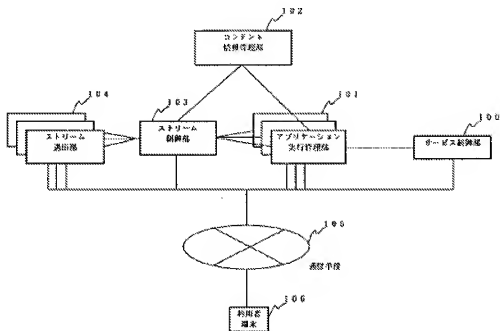
【図14】本発明に係る一般アプリケーション実行管理部の構成管理を示すシーケンス図である。

【符号の説明】

14

- 100 サービス制御部
- 101 アプリケーション実行管理部
- 102 コンテント情報管理部
- 103 ストリーム制御部
- 104 ストリーム送出部
- 105 通信手段
- 106 利用端末
- 120 アプリケーション実行状態情報
- 142 プロセス実行状態情報
- 162 アプリケーション利用状況情報
- 180 プライマリアプリケーション実行管理部
- 181 セカンダリアプリケーション実行管理部
- 182 一般アプリケーション実行管理部

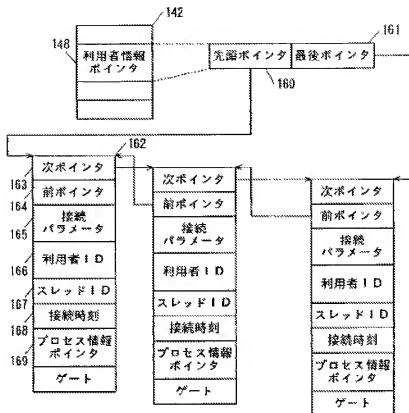
【図1】



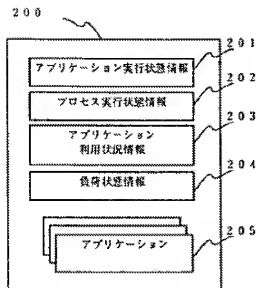
[illegible]

Figure 1 illustrates the structure of process information pointers. It shows a table with three columns: 'アプリケーションID' (Application ID), 'プロセス情報ポインタ' (Process Information Pointer), and 'プロセス情報ポインタの構造' (Structure of Process Information Pointer). The first column is labeled 121, the second 120, and the third 140. The third column is further divided into '先頭ポインタ' (Start Pointer) and '最後ポインタ' (End Pointer). Below this, three vertical lists of pointers are shown, labeled 142, 143, and 144. Each list contains: '次ポインタ' (Next Pointer), '前ポインタ' (Previous Pointer), 'プロセスID' (Process ID), '状態' (Status), '利用者数' (Number of Users), '利用者情報ポインタ' (User Information Pointer), 'アプリケーション情報ポインタ' (Application Information Pointer), and 'ゲート' (Gate).

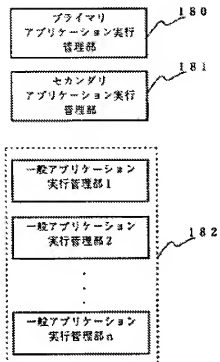
【図4】



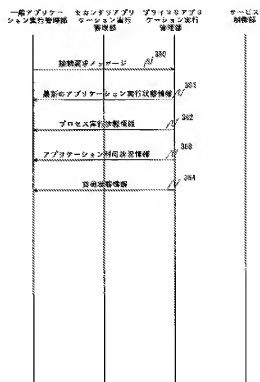
【図6】



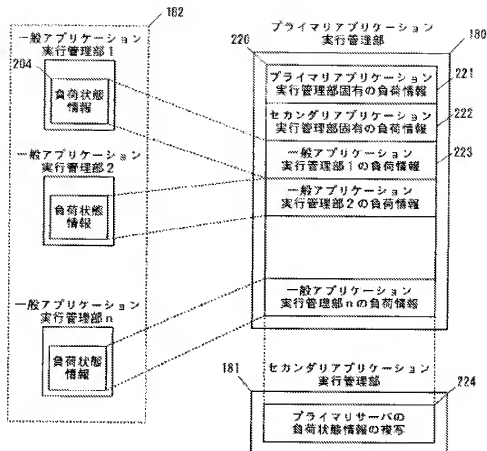
【図5】



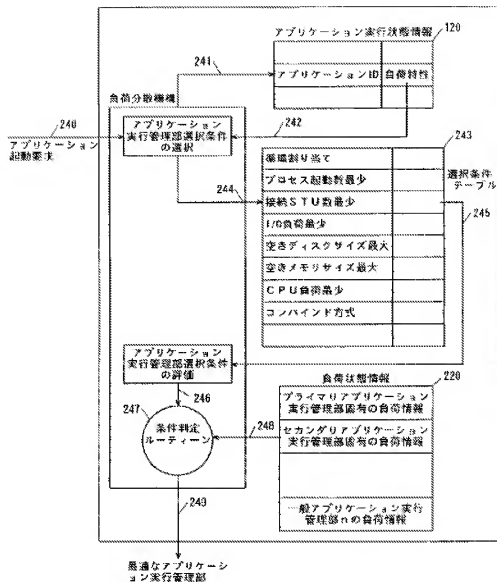
【図14】



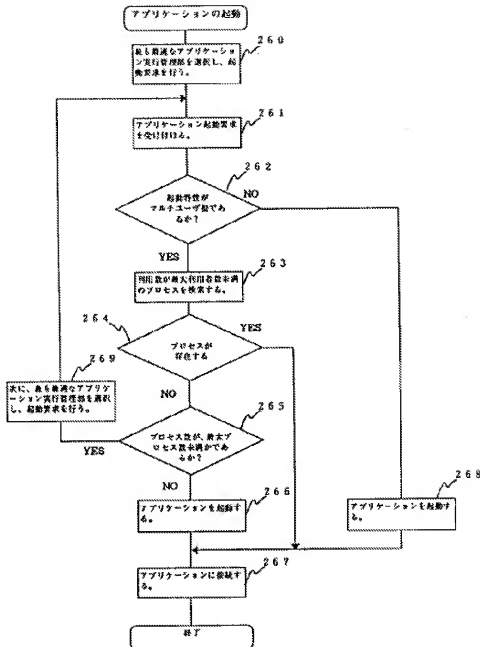
【図7】



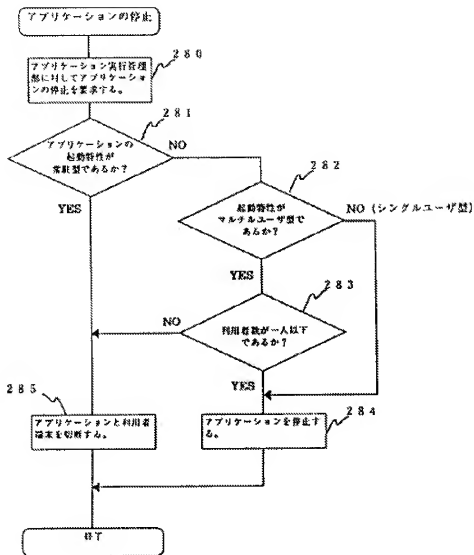
【図8】



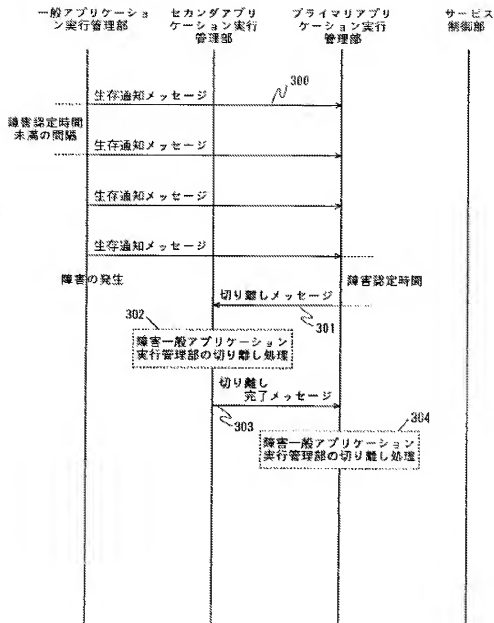
【図9】



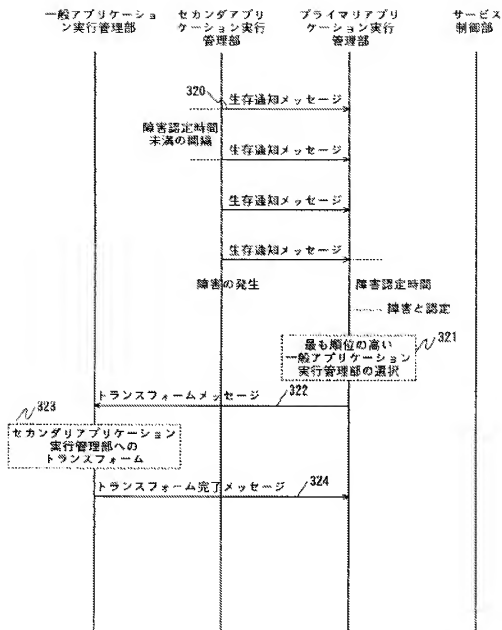
【図10】



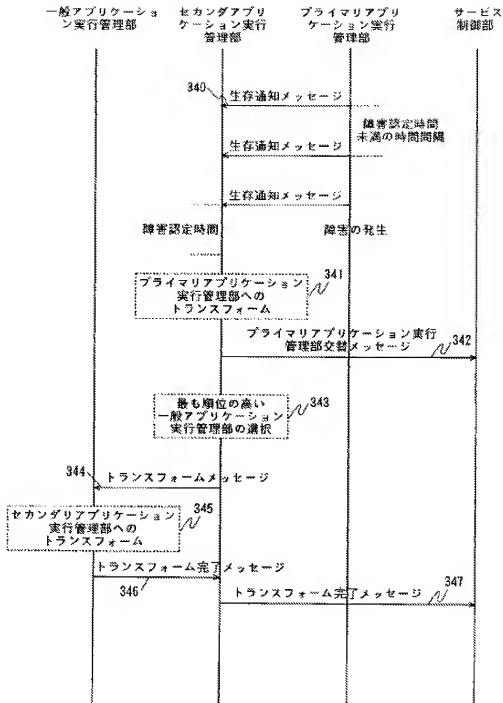
【図11】



【図12】



【図13】



フロントページの続き

(72)発明者 篠原 弘樹
大阪府門真市大字門真1006番地 松下電器
産業株式会社内